DOCUMENT RESUME

ED 305 370 TM 012 837

AUTHOR Ball, Carolyn; And Others
TITLE Development and Testing of the Baseball Coaching Technical Questionnaire (BCTQ).
PUB DATE 86
NOTE 44p.
PUB TYPE Reports - Evaluative/Feasibility (142) -- Reports - Research/Technical (143)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS *Athletic Coaches; *Baseball; Content Validity; Foreign Countries; Item Analysis; *Licensing Examinations (Professions); Models; *Questionnaires; *Test Construction; Test Reliability; Test Validity
IDENTIFIERS *Baseball Coaching Technical Questionnaire; Canada; Coaching; *Paper and Pencil Tests

ABSTRACT

A model and method for developing a valid and reliable paper-and-pencil test for evaluating the technical knowledge of baseball coaches in the National Coaching Certification Program (NCCP) in Canada are discussed. Development of a general theoretical and operational model for the evaluation of NCCP coaches in all sports are also studied. The operational model for test construction includes: (1) selecting a conceptual framework and conducting content analysis with an expert panel; (2) determining uses and details of test administration; (3) identifying the target population; (4) developing items and indicators, including rating scale test specification tables; (5) modifying selected items and test assembly, dimensions, scoring keys, and test procedures; and (6) testing and qualifying to determine reliability and validity. Thirty test items and constructs and 13 demographic and coaching-related variables were selected for the Baseball Coaching Technical Questionnaire (BCTQ). Content validity was assessed via content analysis, item specification, and use of an expert panel. Factor analysis determined the construct validity, and discriminant analysis assessed the criterion-related validity, using 134 certified and non-certified coaches. Discriminant analysis showed that 64.93% of the coaches were correctly classified by the test, with the highest criterion referencing validity for non-certified coaches (72% correctly classified) and Level-II coaches (64.3% correctly classified). Although the BCTQ showed a more than adequate degree of reliability and validity, it should be administered to a larger sample of coaches from other Canadian communities to confirm its national reliability. Five flowcharts and five data tables are provided. (SLD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED305370

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Dick Moriarty

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

DEVELOPMENT AND TESTING OF THE BASEBALL COACHING TECHNICAL QUESTIONNAIRE (BCTQ)

By Carolyn Ball, Dick Moriarty and J. Powell
SIR/CAR University of Windsor
Windsor, Ontario Canada

TM012837

ABSTRACT
Development and Testing of the Baseball Coaching Technical Questionnaire (BCTQ)

By Carolyn Ball, Dick Moriarty and J. Powell
SIR/CAR University of Windsor
Windsor, Ontario Canada

This article reports a model and method for development of a questionnaire to test technical knowledge of baseball coaches exposed through the National Coaching Certification Program (NCCP). The reader is taken through the stages of a conceptual model and the practical steps of an operational model for test construction:

I Selecting a conceptual framework and conducting content analysis with an expert panel;

II Determining uses including types of measure and determining who will administer the tests;

III Identification of target population for application;

IV Tending to items and indicators including item generation, item construction format, formation of items and rating scale test specification tables for selection;

V Modifying selected items and test assembly including dimensions, scoring key and test procedures;

VI Testing and quantification to determine reliability and validity.

Testing of the instrument included analysis to determine validity and reliability. Content validity was determined through content analysis, item specification and use of an expert panel. Factor analysis was used to determine construct validity and discriminant analysis for criterion related validity. Reliability was determined by Cronbach's alpha. Discriminant analysis produced two functions (1. equals trained versus non trained, and 2. equals degree of training). Discriminant analysis resulted in 64.93% of the coaches correctly classified. Non-certified coaches (72%) and level two coaches (64.3%) showed the highest criteron referencing validity while level one coaches showed the lowest (57.1%). Suggestions are included on improvement of validity and reliability of test instruments following an initial application.

This article provides conceptual and operational model and method for development and testing of an instrument related to technical skills.

TM012837

Development and Testing of
The Baseball Coaching Technical Questionnaire (BCTQ)

by

Caroline Ball, Dick Moriarty and Jay Powell

SIR/CAR, University of Windsor

The primary purpose of this study was the development of a valid and reliable paper-and-pencil test instrument for evaluating the technical knowledge of baseball coaches in the National Coaching Certification Program (NCCP).

A secondary purpose of the study involved the development of a general theoretical and operational model of test development for the (technical) evaluation of coaches in all sports within the National Coaching Certification Program.

Practical Need for the Study

At present, there is no method of standardized measurement which has been developed to measure the technical knowledge of baseball coaches within the Level II baseball component (NCCP), or at any other coaching level in the NCCP.

TM012837

A second need for the study concerns the degree to which the content of the Level II and Level III Baseball Canada: Advanced Coaching Manual (MacKenzie, 1984) for baseball is translated into actual clinic settings. It is important to know if the instructional objectives outlined in this manual are being interpreted and understood by coaches in the clinical setting.

There is a need to determine whether the content of the Level II Technical component differs significantly from that of the Level I Technical component in baseball in terms of the retention and use of that information by baseball coaches who have taken the Level II course.

There were two research objectives to be achieved:

1. To determine if the Baseball Coaching Technical Questionnaire (BCTQ) knowledge scale provides a valid measure of NCCP Level II Technical baseball coaching knowledge.

2. To determine if the Baseball Coaching Technical Questionnaire (BCTQ) knowledge scale is a reliable measure of NCCP Level II Technical baseball coaching knowledge.

Measurement of validity, the first research objective, revolved around the following research questions and hypotheses:

1.) How well does the BCTQ reflect the content and objectives of the NCCP Level II Technical baseball course, as contained in resources used in the delivery of the NCCP Level II Technical baseball course (i.e., Baseball Canada: Advanced Coaching Manual and the "Course Conductor's Guide")? Will there be a significant difference between the item content of the BCTQ and the instructional content of the Level II Technical Course in baseball? This is a question of content validity.

2.) What underlying dimensions are measured by the BCTQ, as identified by factor analysis? This is a question of construct validity.

3.) Does the BCTQ perform effectively in correctly classifying cases into predicted groups of baseball technical coaching expertise (i.e., Non-certified, Level I, and Level II baseball coaches)? This is a question of criterion-referenced validity.

4.) Is there a significant difference in the ability of the BCTQ to discriminate among each of the three levels of technical coaching knowledge in baseball (i.e., Non-certified, Level I, and Level II baseball coaches)? This is a criterion-referenced validity question and leads to the following hypothesis:

Ho1: There will be no significant difference in the ability of the BCTQ to discriminate among each of the three levels of technical coaching knowledge in baseball ($p < .05$).

Ha1: There will be a significant difference in the ability of the BCTQ to discriminate among each of the three levels of technical coaching knowledge in baseball ($p < .05$).

5.) Is there a significant difference in the ability of the 30 items and 13 demographic variables in the BCTQ to discriminate among each of the three levels of technical coaching knowledge in baseball (i.e., Non-certified, Level I, and Level II baseball coaches)? This question results in the following criterion-referenced hypothesis:

Ho2: There will be no significant difference in the ability of the 30 test items and 13 variables in the BCTQ to discriminate between levels of technical coaching knowledge in baseball ($p < .05$).

Ha2: There will be a significant difference in the ability of the 30 test items and 13 variables in the BCTQ to discriminate between levels of technical coaching knowledge in baseball ($p < .05$).

Measurement of reliability, the second research objective, focused on the following two research questions:

1.) How reliable is the BCTQ in measuring factors of baseball technical coaching knowledge in the NCCP?

2.) What test length provides a reliable estimate of baseball technical coaching knowledge in the BCTQ?

Review of Related Literature

Baseball Canada (1984) states that the Level II Technical component focuses on coaches involved with athletes who may be beginning a more technically structured and disciplined approach to sport participation. Level II Technical introduces coaches to more advanced technical material, involving strategy and complex skill development. In addition, the Level II Technical component is a prerequisite for entry into the Level III Technical component, which is, at present, under development (Baseball Canada, 1984).

The Educational Process

Brennan (1975) also theorizes that it is useful and practical to view an educational system as consisting of a discrete number of 'objective-related modules', referring to all of those factors in an instructional system that are directly related to a particular instructional objective (Brennan, 1975).

Lewin (1963) postulated that certain psychological-environmental factors may also be included as an estimate of the weight of non-manipulated factors that are associated with learning outcomes. That is, learning is a function of individual aptitude (A) instructional environment (T) and psychological environment (E) and can be conceptualized as L=f(A,T,E) (Lewin, 1963).

Cooley and Leinhardt (1975) focused on the relationship between instructional practices and school performance. The criterion predicted included both academic achievement and attitudes toward school, peers, and instructors. School performance was found to be a function of the constructs of initial abilities, opportunity, motivators, structure, and

instructional events (Haertel, 1985). In the form of an equation, the function is P=f(A,O,M,S,I), where P is the function of school performance, A is the initial ability of the student, O represents opportunities for learning, M constitutes motivators for learning, S is the educational structure, and I represents one or a series of instructional events occurring in the educational process.

Bloom (1976) posits that when quality of instruction is improved, both cognitive and affective entry characteristics (i.e., characteristics which are present prior to training or the introduction of a learning unit) are improved. This is illustrated in Figure 1 - Analysis of Learning (Bloom, 1976) below.

Insert Figure 1 about here

Test Construction

Glaser (1963) suggested that the results of an achievement test can be used to rank individuals in an instructional unit along a relative continuum of ability with respect to a specified group of persons (norm-referencing) or along an absolute continuum of proficiency with respect to a specified behavioural criterion (criterion-referencing) (Brennan, 1975).

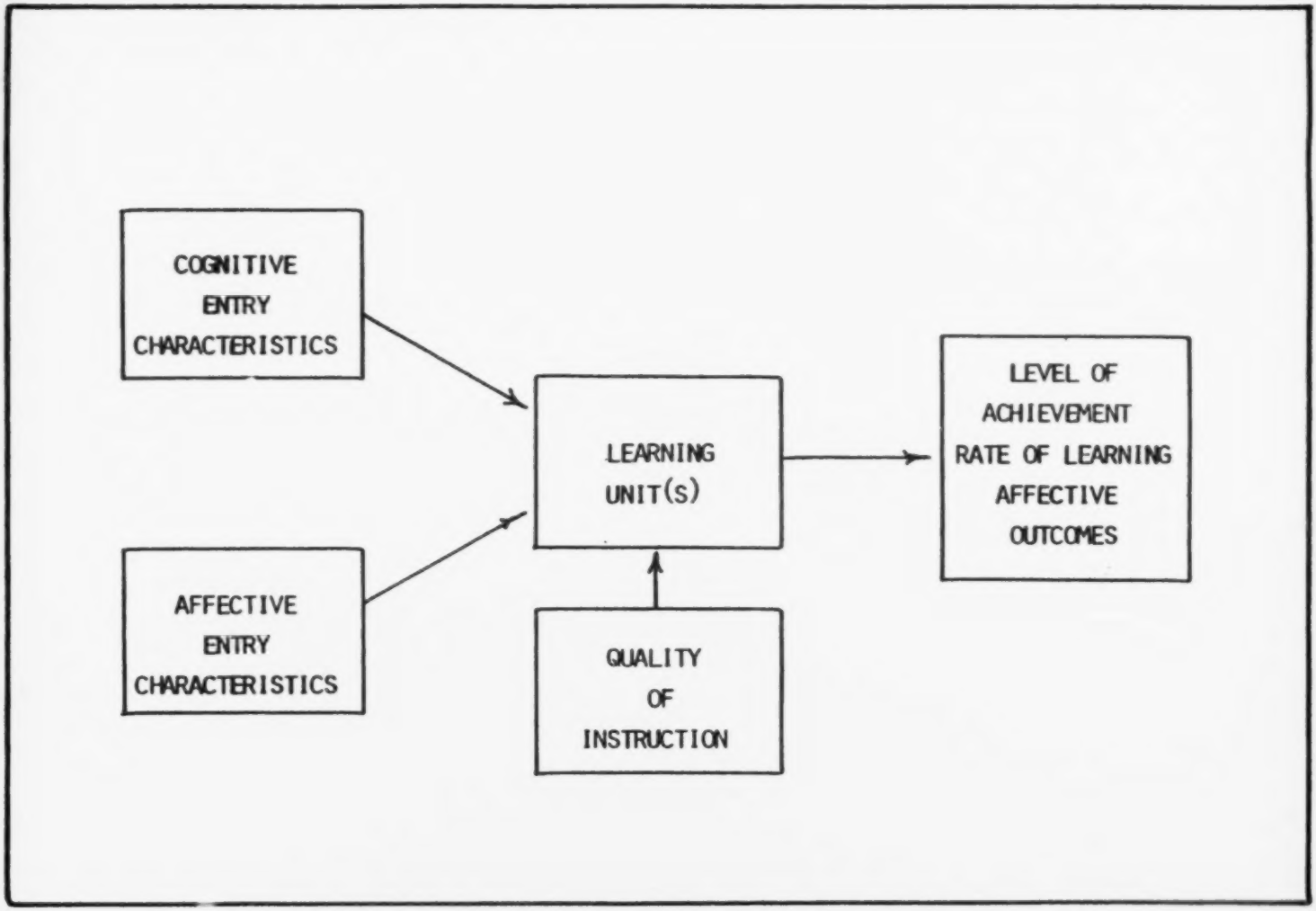


FIGURE 1
ANALYSIS OF LEARNING
(BLOOM, 1976)

Glaser also notes that achievement test scores provide two kinds of information. The first is the degree to which the student has attained criterion performance and the second involves a relative ordering of individuals with respect to their test performance (Brennan, 1975).

Bloom (1974) suggested that if a second learning unit follows the first and the units are sequential (as with the NCCP), that is, the first unit is a prerequisite to the second unit, it will be found that at the end of group instruction some students will have attained mastery of the second unit while others may be in need of more time and help (Bloom, 1974).

In a study by Roid and Haladyna (1982), an 'effective' criterion-referenced test is described as being built from items which logically reflect the intent of instruction. The test must be made up of items that cover all important aspects of the instructional domain upon which the test is based (Roid and Haladyna, 1982).

Green and Lewis (1986) adapted a conceptual overview model for test development procedures, originally devised by Lindeman (1976a, 1976b). This

model is based upon seven stages of test development, including: selecting a conceptual framework (Stage I), determining the use of the instrument (Stage II), specifying the target population (Stage III), paying attention to items or indicators (Stage IV), quantifying selected items (Stage V), testing reliability (Stage VI), and testing validity (Stage VII) (Green and Lewis, 1986).

Insert Figure 2 about here

Construction of the Baseball Coaching Technical Questionnaire (BCTQ)

An operational model was developed by the authors as a means for guiding the construction of the test instrument in a step-by-step process.

Insert Figure 3 about here

Step 1: Content Analysis of the Level II Technical Baseball Program

The primary reference used in developing the Baseball Coaching Technical Questionnaire (BCTQ) was the Baseball Canada: Advanced Coaching Manual (MacKenzie, 1984), which is the prescribed written material for the Level II Technical Baseball Program.

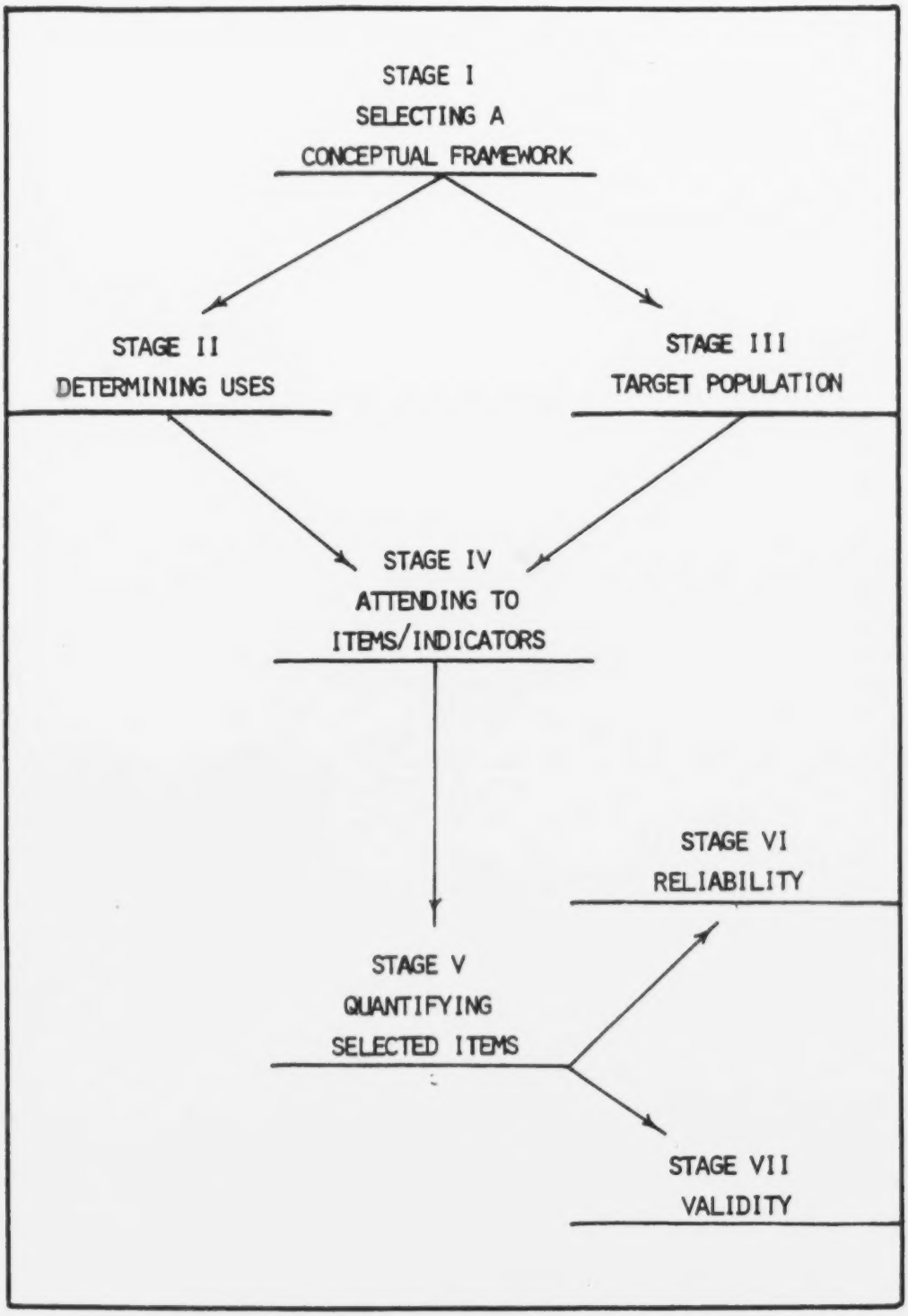


FIGURE 2

CONCEPTUAL MODEL OF TEST CONSTRUCTION
(GREEN AND LEWIS, 1986; ADAPTED FROM LINDEMAN, 1976A, 1976B)

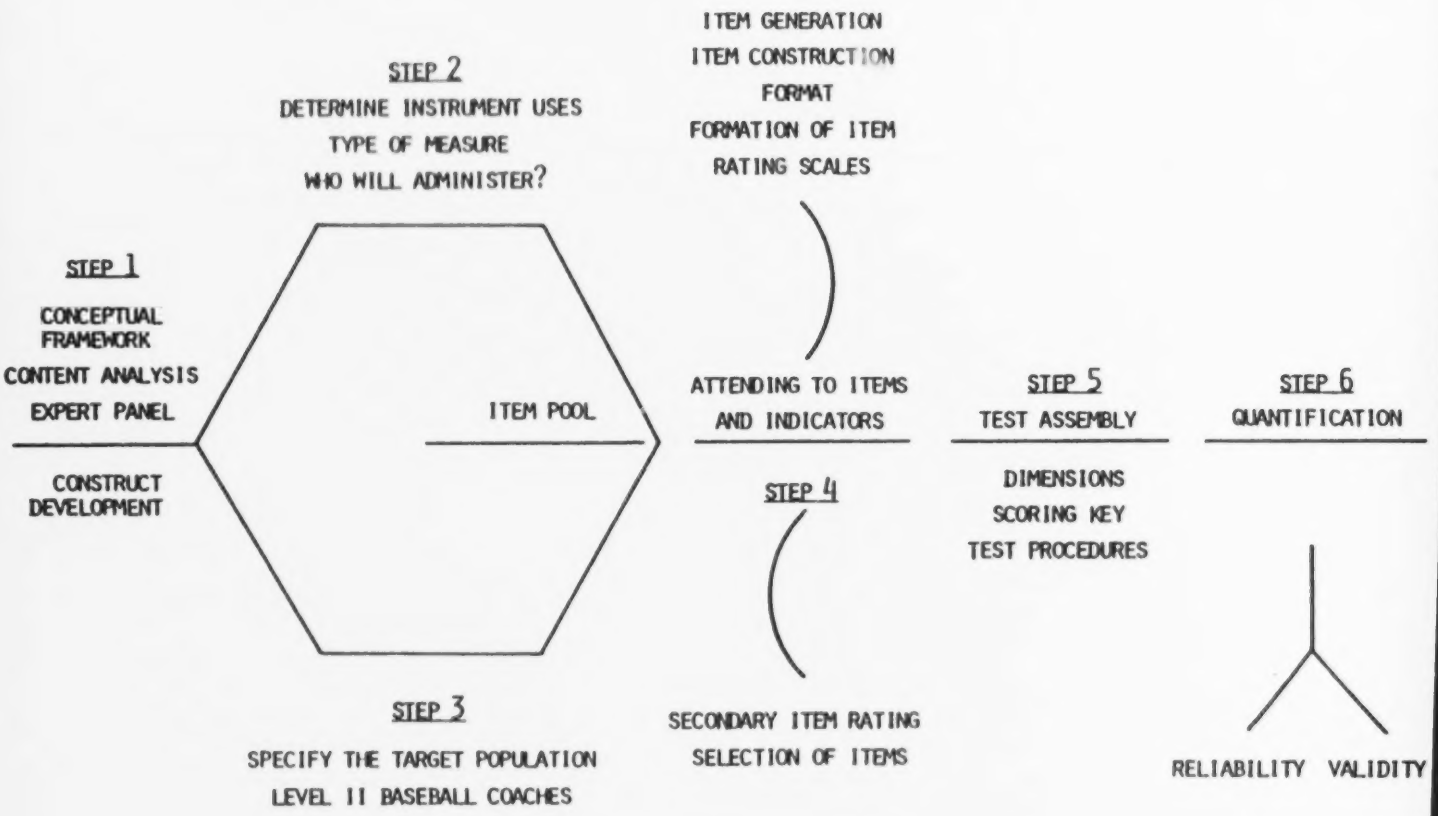


FIGURE 3
OPERATIONAL MODEL - DEVELOPMENT OF THE BCTQ

Content analysis of the Baseball Canada: Advanced Coaching Manual revealed twelve underlying dimensions, including: infielding, pitching, batting, catching, offensive strategy, defensive strategy, baserunning, outfielding, practice sessions, coaching philosophy, physical conditioning, and facilities. The principal researchers working with a panel of 11 experts from the Canadian Federation of Amateur Baseball identified these twelve dimensions utilizing a Conceptual Model of the Level II Technical Baseball Program (See Appendix A), and a table of Test Specifications (See Appendix B).

Step 2: Determining the Use of the BCTQ

The BCTQ has been designed to serve a number of purposes. The total test score represents an overall assessment of baseball, in terms of baseball knowledge, which is relevant to the Level II Technical course and relevant to other coaches at the same level.

The instrument also functions as a method for placement of baseball coaches without any past coaching training. Coaches who do have coaching experience in baseball can be placed into an appropriate level based on set criteria for a specified coaching level.

The BCTQ functions as a method for verification of placement within a NCCP Technical coaching level in baseball. By considering each set of test data individually, it is possible to provide specific technical feedback to the coach.

A final use is that item scores for the group of examinees can serve a diagnostic role.

A simple paper and pencil test was desired which could be administered individually or in a group setting with little or no intervention by the tester, thereby avoiding bias and retaining objectivity.

Step 3: Specifying the Target Population

Two target populations were considered in the development of the BCTQ, the immediate study population and the eventual intended users of the BCTQ. The study population required a heterogeneous sample of coaches from throughout Ontario representative of non-certified and certified coaches at the various levels.

Selection of Participants

A large pool of potential study participants was made possible through the following means:

1. An updated listing of Level I and Level II baseball coaches from the Ontario region (Ontario Baseball Association, 1985). This list is renewed yearly and provides details on all NCCP courses which have been completed by the coach.

2. Updated community association coaching lists, including the level of age coached and addresses. These lists were the primary source for non-certified baseball coaches.

Including both classroom and mailed-in participants in the study, the overall sample size was N=142.

Step 4: Attending to Items and Indicators - Developing BCTQ Items

In May, 1985 a NCCP Level III Technical pilot clinic for baseball was held in Oshawa, Ontario, with a number of baseball coaches and instructors from across Canada. From this group, a panel of 11 raters was recruited by the author to participate as a means for assessing the relative difficulty, complexity, and discriminability of a large pool of test items.

Items were constructed, based on the degree of complexity and difficulty, as statements which were to receive responses on an agree-undecided-disagree scale.

Item content was derived directly from the content of the Baseball Canada: Advanced Coaching Manual (MacKenzie, 1984).

A total of N=104 potential test items were constructed, based on the contents of the Baseball Canada: Advanced Coaching Manual (MacKenzie, 1984) and presented to the panel of baseball experts to be rated on item difficulty and complexity as perceived by the raters.

A table of test specifications (Nitko, 1983; Lien, 1976) was developed to incorporate the cognitive behaviours and construct areas into a detailed graphical representation of potential item areas (See Appendix B).

Panel members were asked to respond to each statement without referring to any related materials. An item analysis was then carried out on each of the responses of the panel members, including the calculation of a difficulty index and a selection ratio for discriminability for each item relative to the Level II Technical course.

Step 5: Test Assembly

Based on difficulty and discriminability, a total of N=30 test items or constructs were chosen for inclusion in the instrument from the stratified sample.

The 30 BCTQ test items were statements considered independent of each other and measured an underlying measure of technical coaching knowledge in baseball. Some test items have parts making a total of N=63 items and parts to be measured. Of the 30 items, N=12 were summated from a multi-part item, that is, they contained either three or four parts, each a partial measure of the same underlying construct.

The Baseball Coaching Technical Questionnaire was constructed in a booklet format, containing three primary sections: 1. directions to the respondent, 2. a NCCP Coach Profile form (CAC, 1984b), and 3. the 30 test items.

For each correct response (i.e., a match) the respondent was to be awarded a score of 1 point. A response which was incorrect (i.e., no match) was subsequently awarded a score of 0. Originally, scores were to be awarded as 1 point for an incorrect

response, 2 points for an undecided response, and 3 points for a correct response. It was felt by the researchers that an indecisive response constitutes a response indicating lack of knowledge and , therefore, responses were re-assigned new values of either 0 or 1 point. Thus, a dichotomous scale of test item responses was created, (agree-disagree).

The dimensions of the BCTQ are such that it is easily distributed i.e., by mail or in a classroom, and easily administered, requiring little prior training.

Analysis of the Data

Since the goal of this research was to develop a written test instrument to increase the construct of technical coaching knowledge in baseball, two research objectives were proposed.

1. To determine if the BCTQ provides a valid measure of NCCP Level II technical coaching knowledge in baseball, and

2. To determine if the BCTQ is reliable in measuring NCCP Level II technical coaching knowledge in baseball.

Insert Figure 4 about here

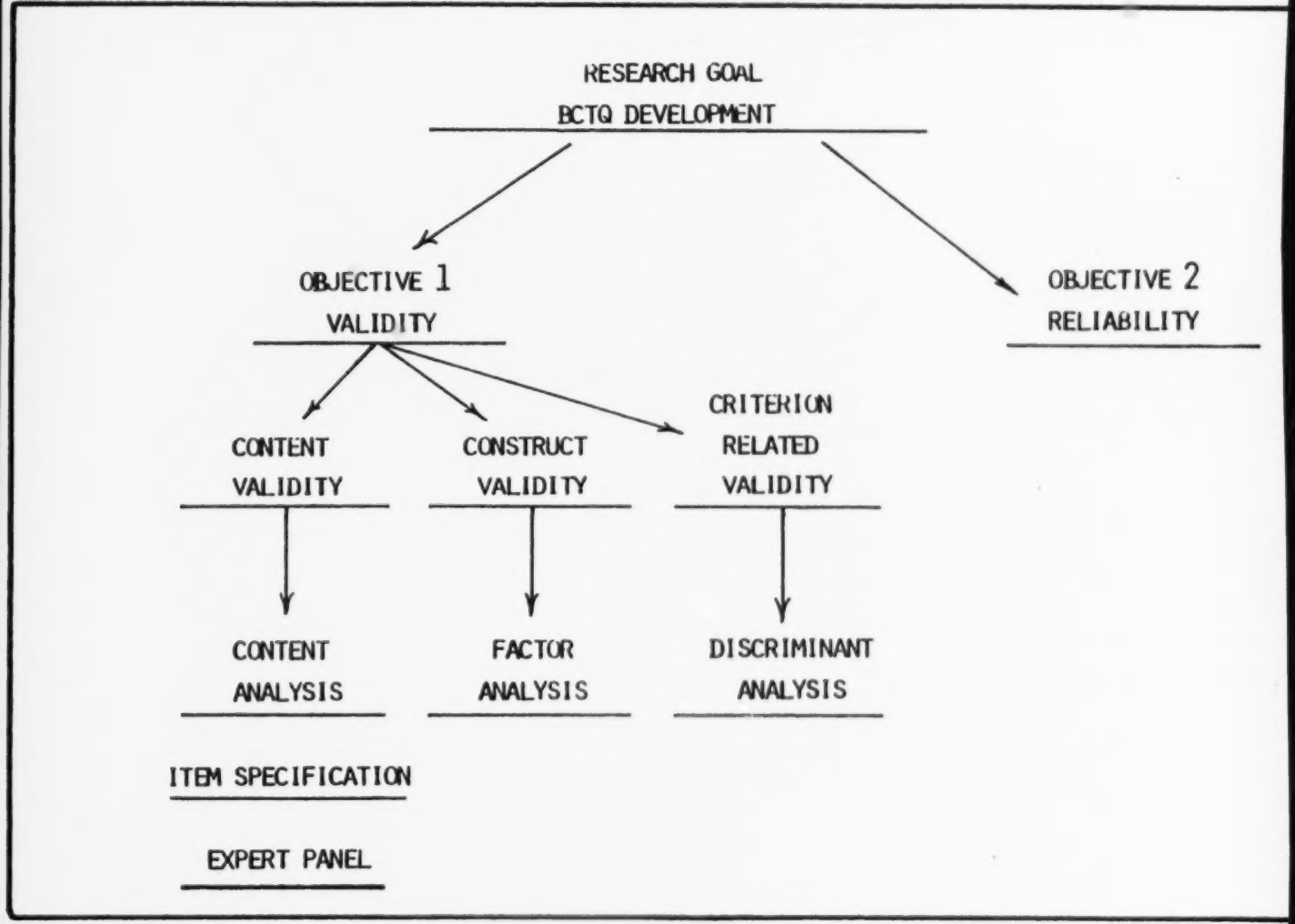


FIGURE 4
MODEL OF RESEARCH OBJECTIVES (BCTQ)

Figure 4 - Model of Research Objectives (BCTQ) illustrates the goals and various statistical techniques utilized to answer these questions.

Construct validity consists of content, construct and criterion related validity. Content validity was determined by content analysis, item specification and expert panel assessment. Factor analysis was used to test construct validity. Discriminate analysis determined criterion related validity. Cronbach's alpha was calculated on the total 63 items contained in the 30 questions to determine reliability. The Statistical Package for the Social Sciences (SPSS) was used as a computer package to assess the data gathered from the BCTQ.

Content Validity

Content validity resulted from an indepth analysis of the Advanced Coaching Manual (CAC, 1984) to determine the type and proportional percentage of content included. A conceptual model and test specifications. A pool of 104 items had been reduced to 30 questions (63 items including subsections) which were selected based on complexity and content as determined by a panel of baseball coaching experts (N=11). Borg (1963) stresses content validity in

test development independent upon comparison of the course content and instructional objectives with the content and objectives of the test instrument.

Table 1 - Test Items in the BCTQ and Initial Statistics Factor Analysis list the areas included in the actual BCTQ and it is proportional to the percentage of content in the Baseball Canada: Advanced Coaching Manual. Scale items with an "I" prefix are indicative of single scale responses to positional statements while a "Q" prefix indicates that the positional statement has a composite of response scales.

Insert Table 1 about here

Nulty (1967) indicates content validity of a test measure should have agreement by potential users of the test, or at least by persons in positions of responsibility (i.e., a panel of experts) involved in a sound and well carried out program, thereby assuring content face validity.

Construct Validity

Cronbach (1971) states that construct validity can be evaluated by investigating the psychological qualities that a test measures. Thus, by means of a factor analysis, it was possible to determine the

	VARIABLE	COMMUNALITY	FACTOR	EIGENVALUE	% VARIANCE	CUMULATIVE %
BASERUNNING	I1	.22989	1	2.90256	9.7	9.7
PHYSICAL CONDITIONING	Q2	.25287	2	2.20155	7.3	17.0
FACILITIES	Q3	.38328	3	1.79005	6.0	23.0
BATTING	I4	.25381	4	1.73010	5.8	28.7
INFIELDING	I5	.21852	5	1.65308	5.5	34.3
PHILOSOPHY	I6	.27630	6	1.47503	4.9	39.2
BATTING	Q7	.29497	7	1.42153	4.7	43.9
BATTING	Q8	.43965	8	1.32844	4.4	48.3
PITCHING	I9	.23772	9	1.21557	4.1	52.4
OUTFIELDING	Q10	.45593	10	1.15658	3.9	56.2
DEFENSIVE STRATEGY	I11	.28456	11	1.08503	3.6	59.9
INFIELDING	I12	.27272	12	1.06149	3.5	63.4
PITCHING	I13	.32509	13	1.01076	3.4	66.8
PITCHING	I14	.22347	14	.93356	3.1	69.9
BATTING	Q15	.37248	15	.91416	3.0	72.9
DEFENSIVE STRATEGY	I16	.22626	16	.88216	2.9	75.9
OFFENSIVE STRATEGY	Q17	.39449	17	.79912	2.7	78.5
PRACTICE SESSIONS	Q18	.28922	18	.79010	2.6	81.2
DEFENSIVE STRATEGY	Q19	.30099	19	.77635	2.6	83.8
PITCHING	I20	.27863	20	.65195	2.2	85.9
BATTING	I21	.32323	21	.57417	1.9	87.9
INFIELDING	I22	.26478	22	.53493	1.8	89.6
CATCHING	I23	.31260	23	.49866	1.7	91.3
BASERUNNING	Q24	.39572	24	.46471	1.5	92.8
OFFENSIVE STRATEGY	I25	.17159	25	.44778	1.5	94.3
INFIELDING	I26	.22175	26	.39697	1.3	95.7
OUTFIELDING	Q27	.19549	27	.38483	1.3	96.9
PITCHING	I28	.19762	28	.33346	1.1	98.1
CATCHING	I29	.31983	29	.30673	1.0	99.1
OUTFIELDING	Q30	.34537	30	.27667	.9	100.0

TABLE I

TEST ITEMS IN THE BCTQ AND
INITIAL STATISTICS - FACTOR ANALYSIS

degree to which certain explanatory concepts or constructs account for performance on a test such as the BCTQ.

An unweighted least squares method (Christofferson, 1975; Muthen, 1978; Muthen, 1981) was used to factor analyze the 30 items on the BCTQ.

Of the total variance, 66.8% was attributable to the first 13 factors, with the remaining 17 factors accounting for only 33.2% of the variance. (See Table 1 - Test Items in the BCTQ and initial Statistics Factor Analysis)

In this study, a limit of 13 factors was determined to be significant as contributing to the variance of the variables.

A comparison of each of the 13 retained factors with those sub-construct content areas outlined in the initial content analysis revealed that the BCTQ does closely approximate the construct of technical coaching knowledge and its related sub-constructs. (See Step 1 Program) Additionally, the extracted factors were matched up with the conceptual model of the Level II Technical course in baseball.

The results of the factor analysis offer evidence, in the form of correlation coefficients, that the construct of technical coaching knowledge in baseball and its related sub-constructs have been adequately measured by the BCTQ.

In terms of criterion-referenced testing, the factors derived from the factor analysis provide evidence that certain variables are more relevant than others in developing criteria for performance based on technical coaching knowledge in baseball. Further testing and analysis of the BCTQ is necessary to confirm the accuracy of these factor derivations.

Criterion Related Validity

Discriminant analysis was used to further examine the construct validity of the BCTQ, as well as to determine its criterion-related validity. A three group discriminant analysis was carried out to distinguish between the three different levels of coaching knowledge, based on the 30 test items and 13 specified demographic and coaching related variables in the BCTQ.

The overall percentage of cases correctly classified was 64.93%. Eighty seven cases were correctly classified into their predicted groups. (See Table II - Classification Results - BCTQ).

Insert Table II about here

Non-Certified Coaches: A total of N=36 (72%) coaches were correctly classified, with an additional N=6 (12%) misclassified as Level I coaches, and N=8 (16%) misclassified as Level II baseball coaches. The moderately high percentage of correct classifications indicates that close to 75% of coaches with non-certified status were classified as having no formal NCCP training in their backgrounds. Coaches misclassified as non-certified indicate that even though these coaches may not have taken a NCCP course of instruction in baseball, their level of technical coaching knowledge may be sufficient to place them in one of these two groups of coaching expertise.

Level I Coaches: Level I coaches had the lowest rate of correct classifications, with only N=24 (57.1%) correct.

ACTUAL GROUP	NUMBER OF CASES	PREDICTED GROUP MEMBERSHIP 1	2	3
NON-CERTIFIED	50	36 (72%)	6 (12%)	8 (16%)
LEVEL I	42	10 (23.8%)	24 (57.1%)	8 (19%)
LEVEL II	42	6 (14.3%)	9 (21.4%)	27 (64.3%)

PERCENT OF 'GROUPED' CASES CORRECTLY CLASSIFIED: 64.93%

TABLE II
CLASSIFICATION RESULTS (BCTQ)

The researchers posit that the fairly low rate of correct classifications for the Level I group indicate that Level I coaches are comprised of a fairly wide range of baseball coaching expertise, since there are no Level I entry qualifications and no standards or criteria presently set by the NCCP to substantiate the attainment of Level I Technical knowledge in baseball. The moderate percentage (64.3%) of correct classifications for this group (N=43) suggests that a number of these coaches may not have met the criterion technical knowledge level for the course content, as stated in the National Baseball Coaching Program: Course Instructors Manual (CAC, 1984).

In terms of test item construction, the overall correctly classified cases percentage suggests that perhaps the items determined to be the best predictors and/or discriminators for each group were low in numbers and that other more non-discriminating variables were present.

Based upon observation of the absolute differences between group means for the canonical discriminant functions, the following labels were given to each of the two discriminant functions:

Insert Table III about here

The researchers posit that the fairly low rate of correct classifications for the Level I group indicate that Level I coaches are comprised of a fairly wide range of baseball coaching expertise, since there are no Level I entry qualifications and no standards or criteria presently set by the NCCP to substantiate the attainment of Level I Technical knowledge in baseball. The moderate percentage (64.3%) of correct classifications for this group (N=43) suggests that a number of these coaches may not have met the criterion technical knowledge level for the course content, as stated in the National Baseball Coaching Program: Course Instsructors Manual (CAC, 1984).

In terms of test item construction, the overall correctly classified cases percentage suggests that perhaps the items determined to be the best predictors and/or discriminators for each group were low in numbers and that other more non-discriminating variables were present.

Based upon observation of the absolute differences between group means for the canonical discriminant functions, the following labels were given to each of the two discriminant functions:

Insert Table III about here

FUNCTION	EIGENVALUE	% VARIANCE	% CUMULATIVE	CANONICAL CORRELATION	WILKS' LAMBDA	X^2	DEGREES OF FREEDOM	SIGNIFICANCE
1	0.46907	62.31	62.31	0.5650658	0.5302651	77.077	40	0.0004
2	0.28370	37.69	100.00	0.4701075	0.7789990	30.344	19	0.0476

TABLE III
CANONICAL DISCRIMINANT FUNCTIONS (BCTQ)

Function 1 = Training vs. Non-Training in Baseball Coaches

Function 2 = Degree of Training in Baseball Coaches

Insert Table IV about here

The null hypothesis (Ho1) of no difference relating to the ability of the BCTQ to discriminate between each of the three coaching levels at the .05 level of confidence was rejected and the alternate (Ha1) considered, based on the difference in group means for each of the two discriminant functions.

From the canonical discriminant functions it was possible to ascertain that there was a difference in the ability of the BCTQ to discriminate between each of the three coaching levels. Since Function 1 was more significant ($p < .001$) than Function 2 ($p < .05$) it is suggested that the BCTQ discriminates best between non-certified and Level I baseball coaches and moderately well between Level I and Level II baseball coaches.

For Function 1, the largest standardized coefficients were associated with the variables of outfielding knowledge, physical conditioning knowledge, pitching knowledge, and defensive strategy. These

GROUP	FUNCTION 1	FUNCTION 2
1	-0.87468	0.05678
2	0.59179	0.62906
3	0.44951	-0.69665

TABLE IV
CANONICAL DISCRIMINANT FUNCTIONS EVALUATED AT GROUP MEANS
(BCTQ)

predictor variables were considered to be the best predictors of trained vs. non-trained baseball coaches.

Degree of training (Function 2) variables with significantly large standardized coefficients included overall knowledge (total score), outfielding knowledge, sex of the coach, and offensive strategy.

For Function 1 (trained vs. non-trained coaches) non-certified baseball coaches were best discriminated by the variables 'defensive strategy' and 'overall technical knowledge'. Level I baseball coaches were best discriminated by the variables 'outfielding knowledge', 'physical conditioning knowledge', 'pitching knowledge', and 'defensive strategy'.

For Function 2 (degree of training) Level I coaching placement was best predicted by the variables 'offensive strategy' and 'pitching knowledge'. Level II placement was best predicted by 'overall technical knowledge', 'outfielding knowledge', 'sex of the coach', 'physical conditioning knowledge', and 'defensive strategy'.

The ability of the BCTQ to predict with accuracy roughly two out of three times overall, and four out of five times in terms of non-certified coaches is very encouraging.

Reliability

Cronbach's alpha was calculated for the BCTQ as .6506 and was considered a more than adequate estimation for this test length (N=63).

Conclusions and Recommendations

1. Based on content analyses of the Advanced Coaching Manual and the BCTQ, it was concluded that the test instrument has content validity. The items and constructs presented in the BCTQ are directly related to the content of the NCCP Level II Technical course in baseball.

2. As a result of a factor analysis of the 30 test items in the BCTQ, it was concluded that the test instrument has construct validity. The BCTQ more than adequately measures the underlying dimensions associated with the construct of 'technical coaching knowledge in baseball' as it relates to the Level II Technical course in baseball.

3. It was concluded that the test instrument retains a fair degree of criterion-related validity. The BCTQ significantly discriminates between coaching levels, based on two discriminant function

equations (i.e., 'training vs. non-training' and 'degree of training'). However, the magnitude of this discriminability was not considered strong, prompting the need for more items which are good predictors of Level II Technical coaching knowledge in baseball.

4. It was concluded that untrained or non-certified coaches were identified with more accuracy than either of the trained groups of coaches (i.e., Level I and Level II baseball coaches).

5. It was concluded that the reliability of the BCTQ (alpha = .6506) following a single administration of the instrument, was more an adequate estimate for a test length of N=63 test items. In addition, it was concluded that, in order to increase the reliability of the BCTQ (i.e., to alpha = .80), the test length would have to increase by approximately 2.15 times the original BCTQ test length or N=135 test items.

Psychometric Concerns

1. The validity of the BCTQ, in terms of its ability to discriminate strongly between levels of technical coaching expertise should be improved as a means for better predicting group placement.

2. The BCTQ could be lengthened by means of additional 'like' (i.e., similar) positional statements in order to improve on the validity and reliability of the scale. A better alternative is to retain the 20 questions which were entered into the discriminant analysis and replace the 10 non-discriminating questions with items of similar construction, in order to better classify coaches into their respective levels of ability.

3. Although the BCTQ has attained a more than adequate level of validity and reliability, the researcher suggests that the BCTQ should be administered to a larger sample of baseball coaches from other Canadian baseball communities in order to confirm the reliability of the test instrument.

NCCP Research Concerns

Additionally, it is recommended that the NCCP and Baseball Canada could use the development and evaluation of the BCTQ as a model for future evaluation of coaches at other levels of technical coaching expertise (i.e., Level I and Levels III to V in baseball).

2. It is recommended that the NCCP and Baseball Canada consider the operationalization of the BCTQ as a model for test development in the remaining two NCCP components (i.e., Practical and Theoretical components).

3. From a more general perspective, it is recommended that future NCCP research into evaluation for any sport and/or physical activity could make use of the theoretical and operational models used in the development of the BCTQ to develop similar criterion-referenced tests of a sport-specific nature.

4. Future NCCP research could also centre on the development of a battery of technical, theoretical, and practical test instrumentations to be used in a more objective certification process for determining the placement of coaches into appropriate levels of expertise and training.

5. Further research could also be conducted in terms of the accuracy of the course conductors in evaluating coaches with and without the use of the BCTQ. Further technical instructors could agument the BCTQ with observation protocols and/or rating scales used by the technical instructors or other observers.

6. In general, non-certified coaches appear to be characterized by a general lack of technical knowledge (Function 1). This observation suggests that all coaches would benefit from technical training, at least as far as technical knowledge is concerned.

The major unanswered question in this study is "Does the acquisition of technical knowledge improve the quality of coaching?" A study similar to this one which considers average team performance (controlling for age of players, experience in years of coaching, and perhaps league competitive quality) against coach's training level and BCTQ scores, using discriminant function analysis would address this question. A large enough sample to average out the impact of an extraordinarily talented team would be needed.

Having shown that the BCTQ distinguished among those with and without technical knowledge, now it must be shown that the presence of technical knowledge improves coaching quality! If "yes" to this question, the justification for technical training and coaching certification is established beyond question.

Bibliography

Bloom, B.S. (1974). Time and Learning. American Psychologist, Sept., 682-688.

Borg, W.R. (1965). Educational Research: An Introduction, New York: David McKay Co., Inc.

Brennan, R.T. (1975). A Model for the Use of Achievement Data and Time Data in an Instructional System. Instructional Science, 3, 113-116.

Christoffersson, A. (1975). Factor Analysis of Dichotomized Variables. Psychometrika, 40(1), 5-32.

Coaching Association of Canada. (1984). National Baseball Coaching Program: Course Conductor's Manual. Ottawa: Author.

Cronbach, L.J. (1970). Essentials of Psychological Testing, (3rd Ed.). New York: Harper & Row Pub.

Cronbach, L.J., & Meehl, P.E. (1955). Construct Validity in Psychological Tests. Psychological Bulletin, 52, 281-302.

Glaser, R. (1976). Components of a Psychology of Instruction: Toward a Science of Design. Review of Educational Research, 46(1), 1-24.

Green, L.W., & Lewis, F.M. (1986). Measurement and Evaluation in Health Education & Health Promotion. Palo Alto, Calif.: Mayfield Pub., Co.

Haertel, E. (1985). Construct Validity and Criterion Referenced Testing. Review of Educational Research, 55(1), 23-46.

Lewin, K. (1963). Field Theory in Social Science. London: Tavistock.

Lien, A.J. (1976). Measurement and Evaluation of Learning. Dubuque: Wm. C. Brown Co., Pub.

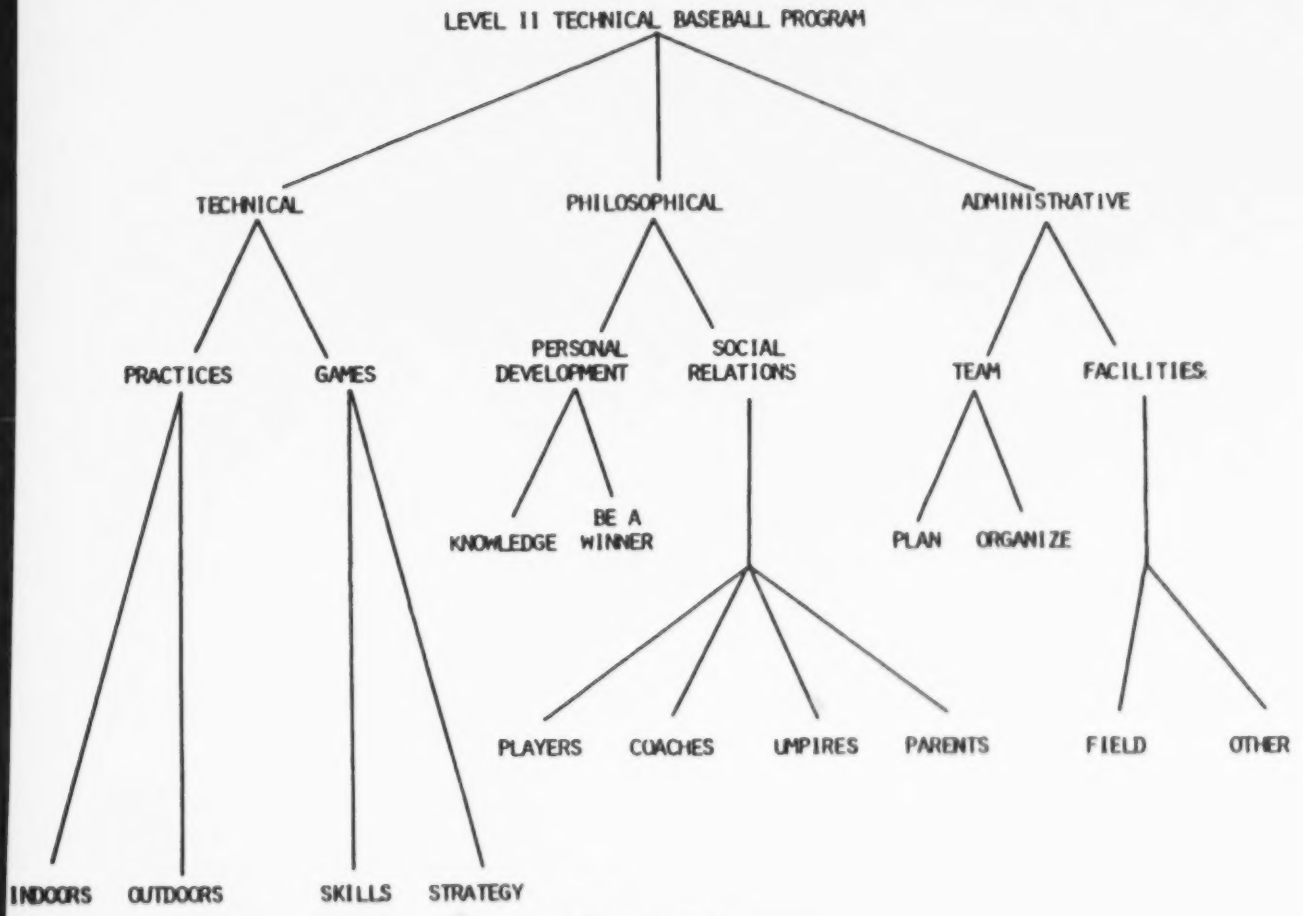
Muthen, B. (1978). Contributions to Factor Analysis of Dichotomous Variables. Psychometrika, 43(4), 551-560.

Muthen, B., & Christoffersson, A. (1981). Simultaneous Factor Analysis of Dichotomous Variables in Social Groups. Psychometrika, 46(4), 407-419.

Nitko, A.J. (1983). Educational Tests and Measurements: An Introduction. New York: Harcourt, Brace, and Yovanovich, Inc.

Nunnally, J.C. (1967). Psychometric Theory. New York: McGraw-Hill Book Co.

Roid, G.H., & Haladyna, T.M. (1982). A Technology for Test-Item Writing. New York: Academic Press



APPENDIX A

CONCEPTUAL MODEL OF THE LEVEL II TECHNICAL BASEBALL PROGRAM

	PHILOSOPHY	BATTING	BASE RUNNING	PITCHING	CATCHING	INFIELDING	OUTFIELDING	DEFENSIVE STRATEGY	OFFENSIVE STRATEGY	PRACTICES	PHYSICAL CONDITIONING	FACILITIES		TOTAL
KNOWLEDGE	1	5	-	2	4	2	4	3	3	2	2	2	--	30
COMPREHENSION	1	3	3	3	2	2	2	2	2	-	-	-	--	20
APPLICATION	1	-	1	5	1	2	-	1	2	-	-	-	--	13
ANALYSIS	-	5	2	2	1	2	2	2	1	-	1	-	--	18
SYNTHESIS	-	1	1	1	2	1	1	1	2	1	-	-	--	11
EVALUATION	1	2	-	3	1	2	-	1	2	-	-	-	--	12
TOTAL	4	16	7	16	11	11	9	10	12	3	3	2	--	104

APPENDIX B

TEST SPECIFICATIONS